ON THE CONCEPTION OF PROBABILITY.

The desire for a mathematically and a philosophically sound introduction to the theory of probability is dictated by the importance of the field of applications, by the fact that the theory of probability is mathematics, and by the philosophical interest attached to the term probability. The applications of the theory were originally confined to problems in gambling, but they are now found in statistics, theory of error, statistical mechanics, insurance, etc. This expansion of the field of usefulness would by itself be a sufficient cause for minor changes in the fundamental conceptions and definitions, just as, for instance, in the steel industry the development in methods and use has been followed by changes in the standard definitions and specifications of iron and steel. To this comes that the introduction of new mathematical and logical methods, such as the axiomatic method, has furnished new view-points for the initial steps in the theory of probability. And the inquiry of philosophy into the nature of probability can never be expected to be answered absolutely and finally. In this way all of the relations of the theory to its applications and sources are likely to exert their influence on the fundamental part of this subject. The fact remains that it is extremely difficult to give a satisfactory definition of probability; quoting Poincaré, it is even hardly possible.[1] These conditions taken together explain why the question "what is probability?" in spite of its long history is still alive, and this serves as an apology for the reappearance of the subject in this paper.

The historical methods of introducing probability may be classified according to their relation to four principal methods. The first three of these will, for the purpose of orientation, be briefly mentioned. The discussion in connection with the fourth of the principal methods forms the main part of this paper. It presents a view-point which is thought to throw some new light on the question. The classification follows:

1. Bayes's definition based on the notion of mathematical expectation (*espérance mathématique, mathematische Erwartung*).

[1] H. Poincaré, *Calcul des probabilités*, ed. 1912, p. 24.

2. The method involving the introduction of the notion of equally possible cases as a fundamental notion.
3. The axiomatic method.
4. The method which will here be called the statistical, which has been used by statisticians, and which is based primarily on the law of great numbers.

1. Bayes's ideas of probability are perhaps the oldest historically. At least according to Von Kries's interpretation, they are underlying, though not definitely formulated, in the classical correspondence between Pascal and Fermat.[2] Bayes expressed those ideas in a definition.[3] An example will explain the principle involved. Playing head and tail with two dollars at stake, with even chance of winning and losing, the value of the chance to win is reasonably estimated as one dollar. Then, according to Bayes's definition, the probability of winning would be expressed as one dollar divided by two dollars, that is 1/2. If in general the reasonable value of the chance—the mathematical expectation—of obtaining A is B then B/A is the probability. It seems evident that the modern student of applied mathematics can hardly expect to obtain the clearest notion of probability by way of Bayes's definition unless he has occupied himself extensively with gambling. It seems that only thereby may one develop the notion of the value of a chance, the notion of the mathematical expectation of gain, as a fundamental conception or as an idea of an existing tangible reality. The definition depends on a reference to this reality. Bayes's definition though interesting has now chiefly historical value.

2. The second principal method, found in many classical discussions, makes use of the expressions equally possible or equally probable. These expressions are taken from the ordinary spoken language, and their meaning remains essentially unchanged after they have been introduced and used in the mathematical theory. Illustrative examples are used among which one is predominant, or fundamental as far as it may be said to represent schematically the formation of any probability. It is the example of the bag containing balls of different color but otherwise the same. Assume p white balls in the bag out of a total of q, then there are p equally possible cases out of q in which one draws a white ball by taking one out. Then by the definition the probability is p/q. This gen-

[2] Von Kries, *Die Prinzipien der Wahrscheinlichkeitsrechnung,* 1886, p. 267.

[3] Th. Bayes, "An Essay toward Solving a Problem in the Doctrine of Chances," *Philosoph. Transactions,* 1763, p. 370.

eral method has its disadvantages. What is "equally possible"? To answer this additional explanations of great length have been deemed necessary at the various times. It is sufficient at this place to mention the classical works by Laplace, J. F. Fries, Lexis, Von Kries, and Bertrand, and besides, two more recent discussions by Lourié and Grelling.[4] In spite of the high value of these works, in spite of the increased understanding of probability due to them, their inevitable extensiveness certainly makes them less accessible than is to be desired. And if the notion of the equally possible cases is maintained as the fundamental idea, it is not unlikely that further discussions on the same basis will be found necessary in the future.

3. The *axiomatic method* has been tried by Broggi and Bohlmann.[5] Terms such as "event A," "probability of event A," "event A independent of B or excluding B," etc., are introduced in the axioms stating the two principal laws of combining probability. The axioms can be stated briefly, and logical rigidity as far as the mathematical theory itself is concerned may thus be obtained. When partially independent notions are to be introduced, such as, for instance, that of continuity of probabilities, then the necessary additional axioms are established without difficulty.[6] Nevertheless, as mentioned by the originators of the axioms, the method solves the questions involved only in part. If one does not know what probability is before the axioms are stated the chance remains that one will not know it afterward either. The logical problem left is essentially that to which the main efforts of the previously mentioned extensive works were devoted. The axiomatic method then has its value as supplementary to other solutions. Between the axioms by themselves and reality there is no bridge.

4. The fourth and last principal method is the *statistical method*

[4] Laplace, *Théorie analytique des probabilités,* philosophical introduction from 2d ed. on. J. F. Fries, *Versuch einer Kritik der Prinzipien der Wahrscheinlichkeitsrechnung,* 1842. W. Lexis, *Zur Theorie der Massenerscheinungen,* 1877. Von Kries, *Die Prinzipien der Wahrscheinlichkeitsrechnung,* 1886. Bertrand, *Calcul des probabilités,* 1889 (introduction). S. Lourié, *Die Prinzipien der Wahrscheinlichkeitsrechnung,* 1910. K. Grelling, *Die philosophischen Grundlagen der Wahrscheinlichkeitsrechnung, Abhandlungen der Friesschen Schule,* 1910.

[5] G. Bohlmann, *Encyc. d. math. Wiss.,* Vol. I, Part II, Art. 1D 4b (1900-1904). U. Broggi, *Die Axiome der Wahrscheinlichkeitsrechnung,* Dissertation, Göttingen, 1907. G. Bohlmann, *Die Grundbegriffe der Wahrscheinlichkeitsrechnung, Atti del 4. Congresso Internazionale dei Matematici,* Roma, 6-11 Aprile 1908, Vol. III, 1909, pp. 244 etc.

[6] The point of view of the continuous probabilities is emphasized by L. Bachelier in his *Calcul des probabilités,* 1912.

based principally on the law of great numbers. Montessus by deriving his definition of "equally possible" from the experience expressed in the law of great numbers becomes a representative of this point of view.[7] The writer believes that the statistical method is best suited to the needs in the present main fields of applications, though it is realized that the completest understanding of the problem is reached by a study of all the historical methods. The discussion which follows will propose a method of obtaining logical rigidity in definitions on the basis of the law of great numbers. We shall in the first place consider only the probability which applies in the theory of probability. Other types of probability, subjective and psychological probabilities, which do not necessarily follow the same laws, will be briefly mentioned afterward.

In order to prepare the way for a definition of the probability in the the theory of probability two preliminary notions will first be introduced: that of a great probability and that of a great number.

First let us consider the expression "great probability." This expression shall first be taken in the sense in which it is used in the ordinary scientific language. It shall express the almost safe, or as good as safe expectation that a certain event will occur. Such great probabilities exist. They are derived essentially from experience, but it is realized that they also contain a subjective element expressed in the decision to believe in a certain regularity that makes predictions possible, or in the decision to disregard certain very slight possibilities as immaterial.

The other preliminary notion is that of the great number. The expression "a great number" shall first be taken in the sense of the ordinary scientific language, but in order to adapt the notion for mathematical use we add the following statement: let the functional forms

$$F_1(N_1, N_2 \ldots . N_k),\ F_2(N_1 \ldots . N_k), \ldots .$$

represent certain uses made of the numbers N_1, $N_2 \ldots . N_k$; let $n_1, n_2, \ldots . n_k$ represent any k numbers less than a certain number n; then N_1, $N_2 \ldots . N_k$ are said to be great numbers with respect to the use F_1, $F_2 \ldots .$, and compared with the number n, when—besides $N_1 \ldots . N_k$ being great numbers in the ordinary sense—the differences

$$F(N_1+n_1, N_2+n_2 \ldots .) - F(N_1, N_2 \ldots .)$$

can be neglected.

[7] Montessus, "La loi des grands nombres," *L'enseignement mathématique*, 1905, pp. 122-138. See also his *Calcul des probabilités*, 1908.

That "great numbers" as just defined exist is a matter of experience. When and whether the differences mentioned can be neglected is not merely a mathematical question, but it depends on the empirical realities to which they apply. The explanation "great with respect to a certain use, compared with a certain number" shall always be understood as added whenever the expression "great number" is used in the theory of probability.

After this preparation it is possible to formulate a definition of probability.

Assume that a group of conditions can be indicated under which a certain event may or may not occur. Assume that the nature of this group of conditions allows their repetition any number of times. Among the conditions will be some which limit the knowledge of what actually happens at the individual reproduction of the conditions. Denote by N_2 the unknown number of times in which the event occurs during N_1 repetitions of the conditions. Assume further that N_1 and N_2 are great numbers with respect to any possible use of the fraction N_2/N_1 as represented by

$$F(N_1, N_2) = \phi(N_2/N_1),$$

N_1 and N_2 thereby being compared with some chosen number n. Assume now that the fraction N_2/N_1 with a great probability can be declared to be equal to some distinct value p, and that this great probability can be made a still greater probability by increasing N_1. Then in the sense of the theory of probability p is the probability of the given event under the given conditions. Briefly expressed, the greatly probable ratio of frequency at a great number of repetitions of the conditions is the probability.

That such "probabilities" exist is a matter of experience. Their existence is identified with the existence of the law of great numbers.

In the philosophical introduction to his Theoretical Physics Volkmann[8] advocates what may be termed a repeated epistemological or knowledge-theoretical cyclus. The present problem allows an application of this method of thought, which, though here it may at first appear so, is not a "vicious circle" but rather a "cyclus of logical convergency."

First note that the conception of probability as defined here depends on the previously defined notion of "great probabilities." But the definition of probability just given allows to consider the great probability as that special case of the general probability in

[8] P. Volkmann, *Einführung in das Studium der theoretischen Physik*, 2d ed., 1913, pp. 349 etc.

which this becomes very nearly equal to one. By adding this consideration a sharpened definition of the term "great probability" is obtained, and again, this improvement in rigidity propagates itself into the definition of the general probability. By re-applying the same method the process of sharpening the definitions may be continued.

It is easy to derive the two fundamental theorems of combining probabilities from the definition given here. The second of the theorems, stating that the probability of the contemporary occurring of two mutually independent events is equal to the product of the probabilities of the single events, requires a special definition of the term independency; such definition can be formulated as follows: the event A is said to be independent of the event B when the ratio of cases in which A occurs at a great number of repetitions of conditions is the same whether the total number of cases or only those cases favorable for the event B are considered.

After these two fundamental theorems Bernoulli's theorem of the great numbers can be derived in the usual way. This theorem throws a new light on the definition of probability and on the notion of great numbers, and thus it opens the way for another application and reapplication of a Volkmann's epistemological cyclus.

We are now ready to discuss briefly other types of probability, namely the psychological and subjective probabilities. These are distinct from the already defined empirical or hypothetically empirical or derived hypothetically empirical probabilities treated in the theory of probability. The psychological probability is the degree of expectation. It expresses itself in certain muscular strains and might be measured through these. Our expectations are not always reasonable or logical, therefore it is evident that they are not subject to the laws of the theory of probability. A type of subjective probability may be defined parallel to the empirical probability as the subjectively expected ratio of frequency at a great number of repetitions of generating conditions. Even the so defined subjective probabilities can only approximately follow the laws of the theory of probability, unless by added axioms they are made dependent on these laws. The understanding of the subjective probability improves the knowledge of the subjective element of the "great probability," one of the terms introduced at the beginning of this development. Here again appears the advantage of the epistemological cyclus.

A final cyclus leads now from the last improved conception of probability to the theory of probability, then to the applications

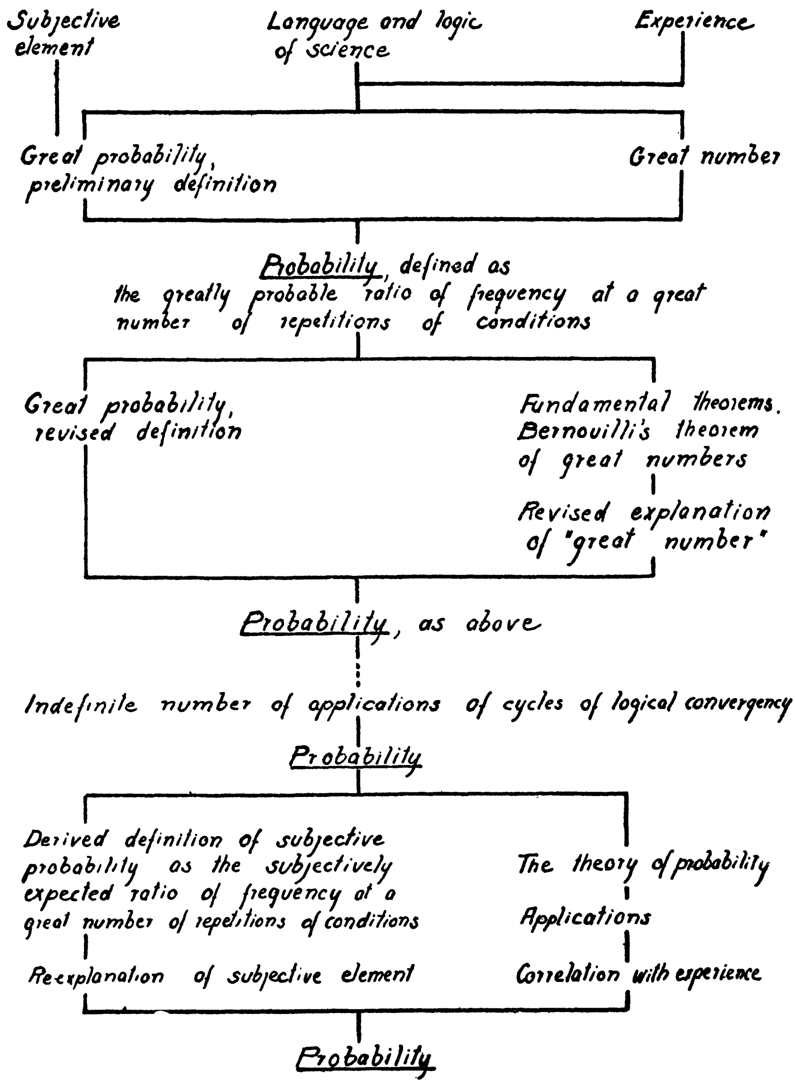

CHART SHOWING PROCESS OF DEFINITION.

of the theory of probability, and therefrom back to the empirical foundation of the definitions.

The chart showing the process of definition reviews the individual steps in this discussion. It should be emphasized that in developing the notion of probability it is necessary to recognize the combined mathematical and empirical nature of the problem. It is conceded that mathematical points, lines, and planes are abstractions. They have neither been observed in the physical world, nor can they be visualized, and thus far, they do not exist outside the paradise of the student of pure mathematics. The relation between the abstract geometrical elements and the corresponding graphical or physical elementary objects is that they are at most mutual approximations. Successful geometry has been developed in spite of that or perhaps on account of that. Abstract probabilities might be derived by eliminating all but the merely mathematical qualities of probability. But in the field of probability the gulf between abstraction and reality is less transparent, and its bridging by proper methods of definition is of decided importance.

H. M. WESTERGAARD.

UNIVERSITY OF ILLINOIS.

RECENT WORK IN MATHEMATICAL LOGIC.

A delightful simplification of the primitive propositions required in Russell's logic (see *Principia Mathematica,* Vol. I, 1910) is made by J. G. P. Nicod in a graceful piece of work (*Proc. Camb. Phil. Soc.,* 1916, XIX, 32-41). It will be remembered that four functions of propositions are used in Russell's logic—not-*p, p* or *q, p* and *q, p* implies *q*: of these, two are taken as indefinables. Nicod makes use of Sheffer's idea (*Trans. Amer. Math. Soc.,* XIV, 481-488) using "*p stroke q*" to mean "not both *p* and *q*" and defines the four functions ordinarily used in terms of this one indefinable. The stroke may be called the sign of *incompatibility*. By means of *three* primitive propositions the propositions required for mathematical logic are developed. The primitive propositions are as follows:

1. If *p* and *q* are elementary propositions, so is *p stroke q*.
2. If *p* and *p stroke* (*r stroke q*) are true, then *q* is true.
3. P *stroke* (π *stroke* Q), where P stands for *p stroke* (*q stroke r*), Q for [*s stroke q*] *stroke* [(*p stroke s*) *stroke* (*p stroke s*)] and π for *t stroke* (*t stroke t*).

The generalized form of the principle of inference (the second